

00862.023126.



PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:)	
	:	Examiner: NOT YET ASSIGNED
HIROAKI IKEDA)	
	:	Group Art Unit: NOT YET ASSIGNED
Application No.: 10/615,304)	
	:	
Filed: July 9, 2003)	
	:	
For: CHARACTER RECOGNITION)	
APPARATUS AND METHOD	:	October 2, 2003

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

SUBMISSION OF PRIORITY DOCUMENT

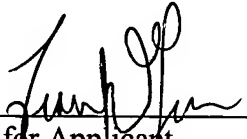
Sir:

In support of Applicant's claim for priority under 35 U.S.C. § 119, enclosed is
a certified copy of the following foreign application:

2002-199505, filed July 9, 2002.

Applicant's undersigned attorney may be reached in our New York office by telephone at (212) 218-2100. All correspondence should continue to be directed to our address given below.

Respectfully submitted,



Attorney for Applicant

Registration No. 42476

FITZPATRICK, CELLA, HARPER & SCINTO
30 Rockefeller Plaza
New York, New York 10112-3801
Facsimile: (212) 218-2200

NY_MAIN 379549v1

日 本 国 特 許 庁
JAPAN PATENT OFFICE

Apple Inc. 10/18/2004
2004-199505

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日 2 0 0 2 年 7 月 9 日
Date of Application:

出 願 番 号 特 願 2 0 0 2 - 1 9 9 5 0 5
Application Number:
[ST. 10/C] : [J P 2 0 0 2 - 1 9 9 5 0 5]

出 願 人 キヤノン株式会社
Applicant(s):

2 0 0 3 年 7 月 2 9 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康 夫



【書類名】 特許願

【整理番号】 4630022

【提出日】 平成14年 7月 9日

【あて先】 特許庁長官殿

【国際特許分類】 G06K 9/42
G06K 9/20
G06K 9/62
G06K 9/68

【発明の名称】 文字認識装置および文字認識方法、プログラムおよび記憶媒体

【請求項の数】 12

【発明者】

【住所又は居所】 東京都大田区下丸子 3 丁目 3 0 番 2 号キャノン株式会社
内

【氏名】 池田 裕章

【特許出願人】

【識別番号】 000001007

【氏名又は名称】 キャノン株式会社

【代理人】

【識別番号】 100079832

【弁理士】

【氏名又は名称】 山本 誠

【選任した代理人】

【識別番号】 100078846

【弁理士】

【氏名又は名称】 大音 康毅

【選任した代理人】

【識別番号】 100087583

【弁理士】

【氏名又は名称】 田中 増顕

【手数料の表示】

【予納台帳番号】 085177

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 0206918

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 文字認識装置および文字認識方法、プログラムおよび記憶媒体

【特許請求の範囲】

【請求項 1】 第 1 の言語で描かれた文字画像を文字認識するのに適した第 1 の認識手段と、

第 2 の言語で描かれた文字画像を文字認識するのに適した第 2 の認識手段と、

前記第 1 の認識手段を用いて文書画像に含まれる複数の文字画像を文字認識した結果に対して、前記第 1 の認識手段で前記第 2 の言語として認識された文字画像を基点として隣接する文字画像の認識結果に基づき、前記第 2 の認識手段を用いて再認識させるための再認識範囲を設定する再認識範囲設定手段と、
を備えることを特徴とする文字認識装置。

【請求項 2】 前記再認識範囲設定手段は、前記再認識範囲に隣接する文字が第 2 の言語であったときに、前記再認識範囲を前記隣接する文字の範囲に拡大することを特徴とする請求項 1 記載の文字認識装置。

【請求項 3】 前記再認識範囲設定手段は、前記再認識範囲に隣接する文字の、前記第 1 の認識手段による認識結果の類似度が所定値以下のときに、前記再認識範囲を前記隣接する文字の範囲に拡大することを特徴とする請求項 1 に記載の文字認識装置。

【請求項 4】 前記再認識範囲設定手段は、各再認識範囲の文字の、前記第 1 の認識手段による認識結果の類似度が所定値以下のとき、当該再認識範囲をそのまま再認識する範囲と決定し、前記類似度が所定値より大きいとき、当該再認識範囲を再認識する範囲から除外することを特徴とする請求項 1 乃至 3 のいずれかに記載の文字認識装置。

【請求項 5】 前記第 1 の認識手段による認識結果と、前記第 2 の認識手段による認識結果とを択一的に選択する認識結果選択手段をさらに備えることを特徴とする請求項 1 乃至 4 のいずれかに記載の文字認識装置。

【請求項 6】 前記認識結果選択手段は、前記再認識範囲の文字の、第 2 の認識手段による認識結果における文字の類似度が所定値以下のときに、第 1 の認識手段の認識結果を選択し、前記類似度が所定値より大きいときに、第 2 の認識

手段の認識結果を選択することを特徴とする請求項 5 に記載の文字認識装置。

【請求項 7】 前記認識結果選択手段は、前記再認識範囲の文字の、第 2 の認識手段による認識結果における類似度の平均値が所定値以下のときに、第 1 の認識手段の認識結果を選択し、前記類似度の平均値が所定値より大きいときに、第 2 の認識手段の認識結果を選択することを特徴とする請求項 5 に記載の文字認識装置。

【請求項 8】 前記認識結果選択手段は、前記再認識範囲の文字の、第 1 の認識手段による認識結果における類似度の平均値と、第 2 の認識手段による認識結果における類似度の平均値とを比較し、前記第 1 の認識手段による類似度の平均値が、第 2 の認識手段による類似度の平均値以下のとき、第 1 の認識手段による認識結果を選択し、前記第 2 の認識手段による類似度の平均値が、第 1 の認識手段による類似度の平均値より大きいとき、第 2 の認識手段による認識結果を選択することを特徴とする請求項 5 に記載の文字認識装置。

【請求項 9】 更に、ネットワークを介して外部装置から前記文書画像を受信する受信手段と、

前記第 1 の認識手段および前記第 2 の認識手段および前記再認識範囲設定手段を用いて、文字認識した結果を前記外部装置に送信する送信手段とを備えることを特徴とする請求項 1 に記載の文字認識装置。

【請求項 10】 第 1 の言語で描かれた文字画像を文字認識するのに適した第 1 の認識手順を用いて、文書画像に含まれる複数の文字画像を文字認識する文字認識ステップと、

前記第 1 の認識手順で第 2 の言語として認識された文字画像を基点として隣接する文字画像の認識結果に基づき、前記第 2 の言語で描かれた文字画像を文字認識するのに適した第 2 の認識手順を用いて再認識させるための再認識範囲を設定する再認識範囲設定ステップと、

前記再認識範囲設定ステップで設定された再認識範囲を、前記第 2 の手順を用いて文字認識する再認識ステップとを備えることを特徴とする文字認識方法。

【請求項 11】 第 1 の言語で描かれた文字画像を文字認識するのに適した

第1の認識手順を用いて、文書画像に含まれる複数の文字画像を文字認識する文字認識ステップと、

前記第1の認識手順で第2の言語として認識された文字画像を基点として隣接する文字画像の認識結果に基づき、前記第2の言語で描かれた文字画像を文字認識するのに適した第2の認識手順を用いて再認識させるための再認識範囲を設定する再認識範囲設定ステップと、

前記再認識範囲設定ステップで設定された再認識範囲を、前記第2の手順を用いて文字認識する再認識ステップと、

をコンピュータに実行させるためのプログラムコードを有する、コンピュータが実行可能なプログラム。

【請求項12】 請求項11に記載のプログラムが格納されたコンピュータ読取可能な記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、原稿を光学的に読み取り文字を認識する文字認識装置、文字認識方法、プログラムおよび記憶媒体に係り、特に英単語を含む日本語文書等、複数の言語を含む文書画像の文字認識が可能な、文字認識装置および文字認識方法、プログラムおよび記憶媒体に関する。

【0002】

【従来の技術】

光学的文字認識装置（OCR）では、濃度射影（ヒストグラム）により文字行を切出し（文字行の抽出）、さらに1文字単位の文字ブロック切出し（文字画像の抽出）を行う。文字ブロックの切出しに際しては、文字行方向に濃度射影を取り、濃度射影値の変化に基づいて文字行を分離し、さらに、各文字行について、文字行と垂直方向に濃度射影を取ることで個々の文字ブロックを抽出する。また1つの文字が複数個の文字ブロックに分離したものについては、標準的な文字サイズや文字ピッチの推定値、および行と垂直方向に濃度射影値等の情報に基づいて、文字ブロックを結合し、1文字単位の文字画像となる、最終的な文字ブロッ

クを生成する。このような文字の切り出し、文字ブロック生成が正しく実行されれば、高精度の文字認識が可能である。

【0003】

しかし、第1の言語に第2の言語が含まれた文書では、正しい文字ブロック生成が行われないことがある。例えば、英単語を含む日本語文書、特に英単語がプロポーションに割り付けされた日本語文書では、英単語部分の文字幅、文字ピッチが不均等であるため、推定値と大幅に異なることがあり、文字ブロックの抽出を誤る場合があり認識精度が低下した。

【0004】

英単語が混在する日本語文書について高い認識精度を得るために、例えば、特開平5-101220号、特開平9-167206号、特開平6-301822号に記載の方法が提案されている。

(1) 特開平5-101220号（従来例1という。）

推定した文字サイズより小さい文字を半角候補文字とし、半角候補文字が連続し且つ半角候補文字列の前または後ろに所定長以上の余白部分が存在した場合、それを英単語候補とし、隣接する1対の半角候補文字のサイズを閾値と比較することにより英単語候補の中に英文字でないと判断される半角候補文字を含む場合は英単語候補から除外し、最終的に英単語候補と判断された半角候補文字は半角英文字であるものとして文書画像からの画像切出が行われ、一方、英単語候補でないと判断された半角候補文字は、隣接する半角候補文字との再結合が行われて切り出される。

(2) 特開平9-167206号（従来例2という。）

一旦、文書画像全体の文字認識を行い、認識結果から英数文字列を抽出し、英数文字列毎にピッチ書式を判定して、そのピッチ書式に応じプロポーションルピッチ用スペース検出処理又は定ピッチ用スペース検出処理を適用し、高精度にスペースを検出する。

(3) 特開平6-301822号（従来例3という。）

空白文字、句読点、括弧等の区切り文字の位置に基づいて、単語としての文字列の照合範囲を定めて、単語辞書との照合を行う後処理を実行する。

【0005】**【発明が解決しようとする課題】**

しかし、従来例 1 では文字ブロック抽出処理は隣接する 1 対の半角候補文字のサイズに基づいて英単語候補かどうか判断して切り出し位置を決定するものであり、例えばプロポーショナルピッチの英単語等で、一部の文字間に接触がある場合は、英単語候補の個々の文字が分離できず、英単語候補を正しく認識できない。また、再認識も行わない。

【0006】

従来例 2 は、英数文字と認識された部分がプロポーショナルかどうか判断するものであり、文字認識処理を誤った場合は英数文字と認識されなかった部分についてはプロポーショナルかどうかの判断すら行わない。また、再認識も行わない。

【0007】

従来例 3 では、区切り文字を用いて単語を抽出するので、区切り文字を認識できなかった場合は単語照合を行うことはできない。

【0008】

本発明は、このような従来の問題点を解消すべく創案されたもので、複数の言語を含む文書画像の文字認識精度を高めることを目的とする。

【0009】**【課題を解決するための手段】**

本発明に係る文字認識装置は、第 1 の言語で描かれた文字画像を文字認識するのに適した第 1 の認識手段と、第 2 の言語で描かれた文字画像を文字認識するのに適した第 2 の認識手段と、前記第 1 の認識手段を用いて文書画像に含まれる複数の文字画像を文字認識した結果に対して、前記第 1 の認識手段で前記第 2 の言語として認識された文字画像を基点として隣接する文字画像の認識結果に基づき、前記第 2 の認識手段を用いて再認識させるための再認識範囲を設定する再認識範囲設定手段とを備える。

【0010】

本発明に係る文字認識方法は、第 1 の言語で描かれた文字画像を文字認識する

のに適した第1の認識手順を用いて、文書画像に含まれる複数の文字画像を文字認識する文字認識ステップと、前記第1の認識手順で第2の言語として認識された文字画像を基点として隣接する文字画像の認識結果に基づき、前記第2の言語で描かれた文字画像を文字認識するのに適した第2の認識手順を用いて再認識させるための再認識範囲を設定する再認識範囲設定ステップと、前記再認識範囲設定ステップで設定された再認識範囲を、前記第2の手順を用いて文字認識する再認識ステップとを備える。

【0011】

本発明に係るコンピュータ実行可能なプログラムは、第1の言語で描かれた文字画像を文字認識するのに適した第1の認識手順を用いて、文書画像に含まれる複数の文字画像を文字認識する文字認識ステップと、前記第1の認識手順で第2の言語として認識された文字画像を基点として隣接する文字画像の認識結果に基づき、前記第2の言語で描かれた文字画像を文字認識するのに適した第2の認識手順を用いて再認識させるための再認識範囲を設定する再認識範囲設定ステップと、前記再認識範囲設定ステップで設定された再認識範囲を、前記第2の手順を用いて文字認識する再認識ステップとをコンピュータに実行させるためのプログラムコードを有する。

【0012】

【発明の実施の形態】

〔文字認識装置〕

図1は、本発明に係る光学的文字認識装置（OCR）等の文字認識装置の一実施形態を示すブロック図である。

【0013】

文字認識装置は、磁気ディスク等の外部記憶装置104およびイメージスキャナ108を備え、原稿の画像データを外部記憶装置104から読み込み、あるいは原稿を直接イメージスキャナ108で読み取る。文字認識装置は全体制御のためのCPU101を備え、CPU101の制御プログラムはROM102に格納されている。イメージスキャナ108から入力された画像はRAM103に保持され、ROM102内の制御プログラムにより、文字認識のための処理を実行す

る。文字認識のための認識辞書は制御プログラムとともにROM102に格納され、あるいは必要な字種ごとに、外部記憶装置104からRAM103に読み込まれる。

【0014】

制御プログラムは複数の言語、例えば日本語、英語に対応しており、それぞれの言語に適した第1の認識手段、第2の認識手段として機能する。

【0015】

なお必要に応じ、文字認識に先立ってノイズ除去等の前処理を実行する。

【0016】

図1において、105はディスプレイ、106はキーボード、107はマウス等のポインティングデバイスである。また、109はネットワークインターフェースであり、図示しない遠隔地に存在する装置と通信し、プログラム、認識辞書、データ等を送受信し得る。

【0017】

次に、文字認識装置において実行される文字認識方法を説明する。

【0018】

[文字認識方法]

[第1の実施形態]

図2、図3は、図1に示す文字認識装置が実行する文字認識方法の第1の実施形態を示すフローチャートである。

【0019】

第1の実施形態では、日本語の認識に適した第1の文字認識手段による文字認識を実行し、認識結果の類似度が低い文字等について、英語の認識に適した第2の文字認識手段による再認識を実行する。

【0020】

図2のフローチャートは、第1の実施形態の全体の流れを示す。

【0021】

ステップS201：まずスキャナ108で読み取った画像あるいは外部記憶装置104に記憶された画像を入力し、ステップS202に移行する。

【0022】

ステップS202：入力された画像を、日本語用の第1の認識手段で文字認識する。ここで行われる文字ブロック抽出は日本語文字を抽出するのに適した文字抽出手法を使用するものとする。文字認識に際しては、従来の技術で述べたように、行方向に射影を取り、行を抽出し、それぞれの行に対して、行とは垂直に射影を取ることで文字ブロックを抽出し、抽出された文字ブロックを用いて文字幅や文字ピッチを推定し、1つの文字が複数のブロックに分離したものについては、推定した文字幅や文字ピッチを用いて1文字になるように結合する。その後、文字ごとに特徴抽出を行い、あらかじめ用意している各字種の参照データと比較し、最も類似度が高い字種を、その文字の認識結果とする。第1の認識手段による日本語用文字認識では、文字ピッチの変動が少なく、かつ文字の縦横比が略正方形であることを想定して、文字ブロックを抽出する。第1の認識手段においても、日本語中で使用されることの多い標準的なフォントのアルファベットは認識可能であるが、対応不能なフォントも多い。画像中のすべての文字について文字認識が終了したら、ステップS203に移行する。

【0023】

ステップS203：ステップS202の認識結果について、各行における類似度が低い文字を含む画像領域を再認識範囲として抽出する。再認識範囲抽出の処理内容については後述する。再認識範囲は、1行以内の画像領域、あるいは複数行にわたる画像領域である。

【0024】

ステップS204：ステップS203において再認識範囲が1つ以上抽出されたか否か判断し、抽出されたときはステップS205に移行する。抽出されなかったときは、そのまま処理を終了する。

【0025】

ステップS205：再認識範囲について、英語用の第2の認識手段によって文字認識する。再認識範囲が複数行にわたるときは、第1の認識手段と同様に行を抽出し、その後文字ブロックを抽出し、文字単位で類似度を調べる。再認識範囲が1行以内の画像領域である場合には、その画像領域から文字ブロックを抽出し

、文字単位で類似度を調べる。ここで行われる文字ブロック抽出は英文字を抽出するのに適した文字抽出手法を使用するものとする。

【0026】

英単語では、文字と文字の接触や、オーバラップ（カーニング）が生じる場合がある。そこで第1の認識手段と異なり、英語用の第2の認識手段では、画素のヒストグラムが少なくなっている所等で文字を分離する。従って、第1の認識手段で分離できなかった文字が分離可能である。また第2の認識手段は、第1の認識手段よりも多様な英語のフォントに対応できるように認識辞書が用意されており、認識精度が向上する。

【0027】

ステップS206：ステップS205で再認識した結果、第1の認識手段よりも高い類似度が得られたときは、その再認識範囲について、第1の認識手段の認識結果を、第2の認識手段の認識結果で置き換える。これによって、英文字が高精度に認識されるようになる。

【0028】

これによって、第1の実施形態は、英単語を含む日本語文書について、日本語文字も英文字も高い認識精度を得ることができる。また、再認識を行う第2の認識手段は第1の認識手段とは独立に動作可能であり、認識手段を日本語や英語以外の異なる言語に適した認識手段に入れ替えることも容易である。従って、第1、第2の認識手段によって多様な言語の組み合わせに対応し得る。

【0029】

次に、ステップS203の再認識範囲の設定について、図3のフローチャートを用いて更に詳細に説明する。

【0030】

ステップS301：閾値T1、T2を設定し、ステップS302に移行する。閾値T1は、再認識範囲を、第1の認識手段で抽出された文字ブロックから前後の文字を含む領域に拡大するか否かの判断に使用する。閾値T2は、設定された再認識範囲について第2の認識手段による認識を行うか否かの判断に使用する。

【0031】

ステップ S 3 0 2：ステップ S 3 0 3 以下の処理の終了判断を行う。図 2 のステップ S 2 0 2 で抽出された全ての文字についてステップ S 3 0 3 以下の処理が完了したときは、そのまま処理を終了する。完了していなかったときはステップ S 3 0 3 に移行する。

【0032】

ステップ S 3 0 3：日本語用の第 1 の認識手段によって、アルファベット、あるいはアルファベットと数字の組み合わせ（以下これらをアルファニューメリックという。）が検出されたか否か判断する。日本語文書に英単語が含まれていなかったときは全ての文字が日本語をして認識され、英単語が含まれていたときにアルファニューメリックが検出される。

【0033】

ステップ S 3 0 4：ステップ S 3 0 3 でアルファニューメリックが検出されたか否か判断し、検出されなかったときは、ステップ S 3 0 2 に戻って未検査の認識結果があるか判断する。アルファニューメリックが検出されたときは、ステップ S 3 0 5 に移行する。ステップ S 3 0 5～S 3 0 8 では再認識範囲を設定する。

【0034】

ステップ S 3 0 5：再認識範囲を後方に拡大する条件について判定する。再認識範囲は、初期状態では、第 1 の認識手段によって検出されたアルファニューメリックの範囲であるが、一定の条件を満たすとき、ステップ S 3 0 6 において再認識範囲を後方の文字に拡大する。再認識範囲を後方に拡大する条件とは、再認識範囲の次の文字がアルファニューメリックであること、もしくは、再認識範囲の次の文字の認識結果の類似度が閾値 T 1 より小さいことである。これらいずれかの条件が満たされたとき、ステップ S 3 0 6 に移行する。再認識範囲の次の文字がいずれの条件も満たさないとき、あるいは次の文字が存在しないときはステップ S 3 0 7 に移行する。なお再認識範囲が 1 行の終端に達していたときは、次の文字は、次の行の最初の文字である。再認識範囲が次の行に拡大されることにより、再認識範囲は複数行にわたることになる。

【0035】

ステップ S 3 0 6 : 再認識範囲の次の文字を再認識範囲に含め、再認識範囲を拡大する。その後、再びステップ S 3 0 5 に戻る。

【0036】

ステップ S 3 0 7 : 初期状態の再認識範囲あるいはステップ S 3 0 6 で後方に拡大された再認識範囲につき、その再認識範囲を前方に拡大する条件について判定する。再認識範囲を前方に拡大する条件とは、再認識範囲の前に存在する文字がアルファニューメリックであること、もしくは、再認識範囲の前の文字の認識結果の類似度が閾値 T 1 より小さいことである。これらいずれかの条件が満されたとき、ステップ S 3 0 8 に移行する。再認識範囲の前の文字がいずれの条件も満たさないとき、あるいは前の文字が存在しないときはステップ S 3 0 9 に移行する。なお再認識範囲が 1 行の始端に達していたときは、前の文字は、前の行の最後の文字である。再認識範囲が前の行に拡大される場合にも、再認識範囲は複数行にわたる。

【0037】

ステップ S 3 0 8 : 再認識範囲の前の文字を再認識範囲に含め、再認識範囲を拡大する。その後、再びステップ S 3 0 7 に戻る。

【0038】

ステップ S 3 0 9 : ステップ S 3 0 5 ~ S 3 0 8 で設定された再認識範囲の文字に、類似度が T 2 以下の文字が存在するか否かを判定する。類似度 T 2 以下の文字が存在すれば、第 1 の認識手段による認識結果の信頼性に問題があると判断し、ステップ S 3 1 1 に移行する。類似度 T 2 以下の文字が存在しないときは、再認識範囲の文字が、アルファニューメリックでありかつ比較的類似度が高い文字であると判断する。これは再認識範囲の文字に対する認識結果が信頼できることを意味するので、ステップ S 3 1 0 に移行する。例えば固定ピッチで書かれたアルファニューメリックは、日本語用文字認識でも十分な精度が得られる場合があり、類似度は T 2 より大となる。一方、可変ピッチで書かれた場合は類似度は T 2 以下となり、誤認識しやすい。

【0039】

ステップ S 3 1 0 : ステップ S 3 0 9 で文字認識結果の類似度が T 2 より大き

いと判断された再認識範囲を、再認識を実行する再認識範囲から除外し、ステップS302に戻る。これによって、英単語であっても、誤認識の可能性の低い範囲は再認識の処理を実行しないので、不要な処理を省略でき、全体の処理速度を高め得る。

【0040】

ステップS311：ステップS309で文字認識結果の類似度がT2以下の文字が存在すると判断された再認識範囲を、再認識を実行する再認識範囲として確定し、ステップS302に戻る。

【0041】

第1の実施形態では、誤認識の可能性が高い英単語の範囲を、再認識範囲として自動的に抽出でき、認識精度を高めることができる。

【0042】

閾値T1、T2は、実験等を行うことにより予め設定された値を用いることとしてもよいし、字種などにより異なる値に設定してもよい。例えば、平仮名、カタカナ、漢字、アルファニューメリック等、注目する認識結果の字種により閾値を使い分けてもよい。

【0043】

第1の実施形態では英単語を含む日本語文書の例を説明したが、他の言語の文書、例えば、中国語文書、韓国語文書、その他の文書であって、英語、ラテン語、その他文字配列が異なる単語を含む文書においても、同様の効果を奏する。

【0044】

[第2の実施形態]

図4、図5は、図1に示す文字認識装置が実行する文字認識方法の第2の実施形態を示すフローチャートである。

【0045】

第2の実施形態では、第1の文字認識手段による認識結果に対して、第2の認識手段による再認識の処理を実行し、さらに、第2の文字認識手段による認識結果に対して、第3の認識手段による再認識の処理を実行する。

【0046】

図4のフローチャートは、第2の実施形態の全体の流れを示す。

【0047】

ステップS401：第1の実施形態のステップS201と同様、画像を入力する。

【0048】

ステップS402：第1の実施形態のステップS202と同様、入力された画像を第1の認識手段で文字認識する。

【0049】

ステップS403：第1の実施形態のステップS203と同様、再認識範囲を設定する。

【0050】

ステップS404：第1の実施形態のステップS204と同様、再認識範囲が1つ以上抽出されたか否か判断する。再認識範囲が抽出されたときはステップS405に移行し、抽出されなかったときは、そのまま処理を終了する。

【0051】

ステップS405：再認識範囲について、第2の認識手段によって文字認識し、文字単位で類似度を調べる。必要に応じて、行抽出、文字ブロック抽出を実行する。

【0052】

ステップS406：第2の認識手段で得られた文字認識結果を採用するかどうか判定する。判定方法については後述する。第2の認識手段で得られた認識結果を採用すると判定したときにはステップS407に移行し、第2の認識手段で得られた認識結果を採用しないと判定したときにはステップS408に移行する。

【0053】

ステップS407：第1の認識手段で得られた文字認識結果のうち、再認識範囲を、第2の認識手段による再認識の認識結果で置き換える。

【0054】

ステップS408：再認識範囲を第3の認識手段で文字認識し、ステップS409に移行する。

【0055】

ステップS409：第3の認識手段による文字認識結果を採用するか否か判定する。判定方法は、ステップS406の判定方法と同様であり、その詳細は後述する。第3の認識手段で得られた認識結果を採用すると判定したときにはステップS410に移行し、第3の認識手段で得られた認識結果を採用しないと判定したときは、そのまま処理を終了する。

【0056】

ステップS410：第1の認識手段で得られた文字認識結果のうち、再認識範囲を、第3の認識手段による再認識の認識結果で置き換える。

【0057】

図5は、ステップS406における判定方法を示すフローチャートである。

【0058】

ステップS501：再認識範囲の、認識結果の類似度の平均値を求める。

【0059】

ステップS502：平均値をあらかじめ定めた閾値T3と比較し、閾値T3以下のとき、ステップS503に移行し、閾値T3より大きいとき、ステップS504に移行する。

【0060】

ステップS503：第2の認識手段による再認識の認識結果を採用せず、第1の認識手段による認識結果をそのまま採用する。

【0061】

ステップS504：第1の認識手段による認識結果を、第2の認識手段による再認識の認識結果によって置き換える。

【0062】

図6は、ステップS409における判定方法を示すフローチャートである。

【0063】

ステップS601：図5のステップS501と同様、類似度の平均値を求める。

。

【0064】

ステップ S 6 0 2：平均値をあらかじめ定めた閾値 T 4 と比較し、閾値 T 4 以下のとき、ステップ S 5 0 3 に移行し、閾値 T 4 より大きいとき、ステップ S 6 0 4 に移行する。

【0065】

ステップ S 6 0 3：第 3 の認識手段による再認識の認識結果を採用せず、第 1 の認識手段による認識結果をそのまま採用する。

【0066】

ステップ S 6 0 4：第 1 の認識手段による認識結果を、第 3 の認識手段による再認識の認識結果によって置き換える。

【0067】

なお図 5、図 6 の処理における閾値 T 3、T 4 は同一であってもよいし、経験則や実験結果等に基づき異なる値に設定してもよい。

【0068】

図 7 は、図 5 に示した判定方法の他の態様を示すフローチャートである。

【0069】

ステップ S 7 0 1：第 1 の認識手段による認識結果の類似度の平均値 M 1 を求める。

【0070】

ステップ S 7 0 2：第 2 の認識手段による認識結果の類似度の平均値 M 2 を求める。

【0071】

ステップ S 7 0 3：平均値 M 1、M 2 を比較し、平均値 M 1 が M 2 以上のときステップ S 7 0 4 に移行し、平均値 M 2 が M 1 より大きいときステップ S 7 0 5 に移行する。

【0072】

ステップ S 7 0 4：第 2 の認識手段による再認識の認識結果を採用せず、第 1 の認識手段による認識結果をそのまま採用する。

【0073】

ステップ S 7 0 5：第 1 の認識手段による認識結果を、第 2 の認識手段による

再認識の認識結果によって置き換える。

【0074】

なお図7の判定方法は第1の認識手段による認識処理と、第3の認識手段による認識処理とを比較、選択する際にも同様にして、判定に適用し得ることはいうまでもない。

【0075】

第2の本実施形態では、類似度の平均値より認識結果の信頼性を判定するので、大局的な認識精度評価が可能である。また複数段階の認識手段を順次適用し、結果を比較するので、多様な言語、字種に対応し得るとともに、複数の認識結果から最適認識結果を選択し得る。

【0076】

[プログラム]

図1の文字認識装置における制御プログラムには、文字認識方法の第1の実施形態および／または第2の実施形態を実行するためのプログラムが含まれる。第1の実施形態を実行するためのプログラムはステップS201～S206、ステップS301～S311の処理を実行するルーチンを含む。第2の実施形態を実行するためのプログラムはステップS401～S410、ステップS501～S504、ステップS601～S604、ステップS701～S705の処理を実行するルーチンを含む。

【0077】

[記憶媒体]

本発明にかかる文字認識方法は図1の文字認識装置以外の装置、例えば汎用的なコンピュータでも実施可能であり、その場合、記憶媒体等で提供される制御プログラムを汎用コンピュータの外部記憶装置にセットし、オペレータ等の指示によりCPU等で実行してもよい。

【0078】

図8は、文字認識方法の第1および第2の実施形態を実行するためのプログラムを記憶した記憶媒体を示すメモリマップである。

【0079】

図 8 においてアドレス A 0 乃至 A 1 にはディレクトリ情報が格納され、A 1 の次のアドレスからアドレス A 2 までの領域には第 1 の実施形態における図 2 のフローチャートを実行するためのプログラムが格納され、A 2 の次のアドレスからアドレス A 3 までの領域には第 1 の実施形態における図 3 のフローチャートを実行するためのプログラムが格納され、A 3 の次のアドレスからアドレス A 4 までの領域には第 2 の実施形態における図 4 のフローチャートを実行するためのプログラムが格納され、A 4 の次のアドレスからアドレス A 5 までの領域には第 2 の実施形態における図 5 のフローチャートを実行するためのプログラムが格納され、A 5 の次のアドレスからアドレス A 6 までの領域には第 2 の実施形態における図 6 のフローチャートを実行するためのプログラムが格納され、A 6 の次のアドレスからアドレス A 7 までの領域には図 7 のフローチャートを実行するためのプログラムが格納されている。

【0080】

特に図示しないが、各種プログラムに従属するデータもディレクトリ情報として管理し、また、各種プログラムをコンピュータにインストールするためのプログラム等をディレクトリ情報内に記憶してもよい。

【0081】

[文字認識システム]

図 9 は、本発明に係る文字認識方法を複数の機器により実行する文字認識システムの一実施形態を示すブロック図である。

【0082】

この実施形態では、インターネット・サービス・プロバイダにおいて、文字認識のための認識辞書を含むプログラムがサーバ S V 1 に格納されている。インターネット・サービス・プロバイダは通信手段を介してインターネットに接続され、インターネットからサーバ S V 1 にアクセスすることによって、複数のユーザが文字認識プログラムを利用し得る。

【0083】

ユーザはスキャナ S C 1、パーソナル・コンピュータ P C 1 を有し、スキャナ S C 1 で読み取った原稿の画像をサーバ S V 1 に送信する。サーバにおいては、

原稿の画像を入力画像として上述した第 1 実施形態又は第 2 実施形態の文字認識の処理を実行し、その認識した結果をパーソナルコンピュータ P C 1 に返信する。

【0084】

なお認識辞書は、必ずしも文字認識プログラムと同一のコンピュータ上に存在する必要はなく、通信手段等を介して参照可能であればよい。例えば、文字認識プログラムをサーバ S V 1、認識辞書を別個のサーバ S V 2 に設け、別個のサービスとすることも可能である。この場合文字認識プログラムの所有者、認識辞書の所有者それぞれがプログラム、辞書という知的資産を内部に保有でき、違法コピー等を防止できる。

【0085】

【発明の効果】

本発明によれば、複数の言語で記載された文書画像の文字認識精度を高め得る。

【図面の簡単な説明】

【図 1】 本発明に係る文字認識装置の一実施形態を示すブロック図である。

【図 2】 図 1 の文字認識装置において実施される文字認識方法の第 1 の実施形態を示すフローチャートである。

【図 3】 図 2 におけるステップ S 2 0 3 の処理を示すフローチャートである。

【図 4】 図 1 の文字認識装置において実施される文字認識方法の第 2 の実施形態を示すフローチャートである。

【図 5】 図 4 におけるステップ S 4 0 6 の処理を示すフローチャートである。

【図 6】 図 4 におけるステップ S 4 0 9 の処理を示すフローチャートである。

【図 7】 図 5、図 6 の処理に替わる処理を示すフローチャートである。

【図 8】 図 1 の文字認識装置で実行される各種データ処理プログラム格納

する記憶媒体のメモリマップである。

【図 9】 本発明に係る文字認識方法を複数の機器により実行する文字認識システムを示すブロック図である。

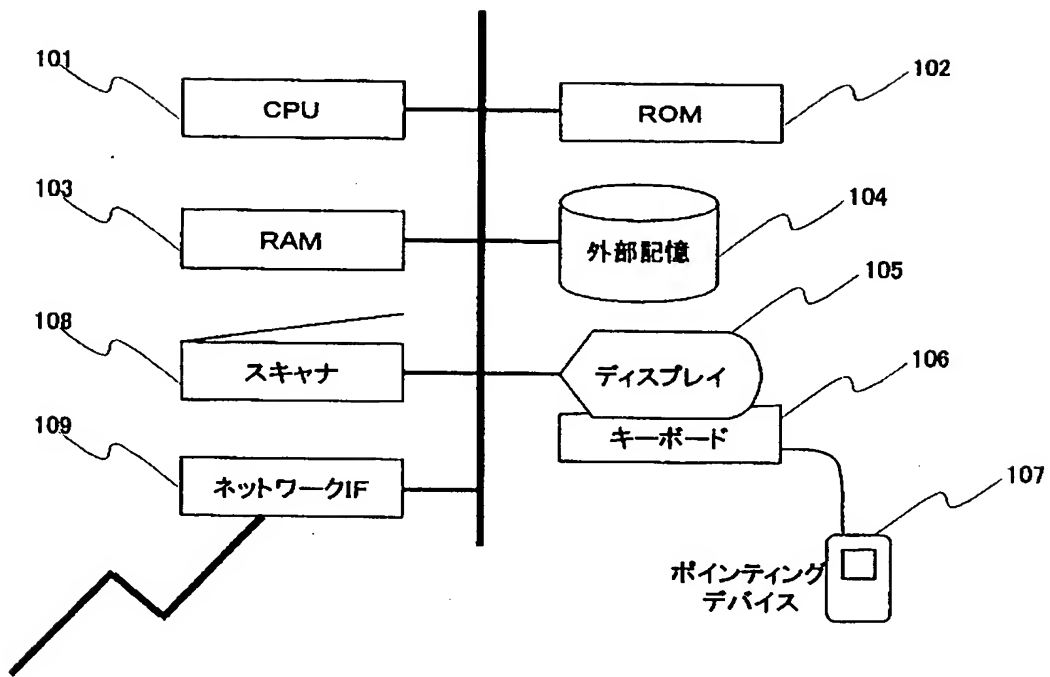
【符号の説明】

1 0 1 . . . C P U
1 0 2 . . . R O M
1 0 3 . . . R A M
1 0 4 . . . 外部記憶装置
1 0 5 . . . ディスプレイ
1 0 6 . . . キーボード
1 0 7 . . . ポインティングデバイス
1 0 8 . . . イメージスキャナ
1 0 9 . . . ネットワークインターフェース

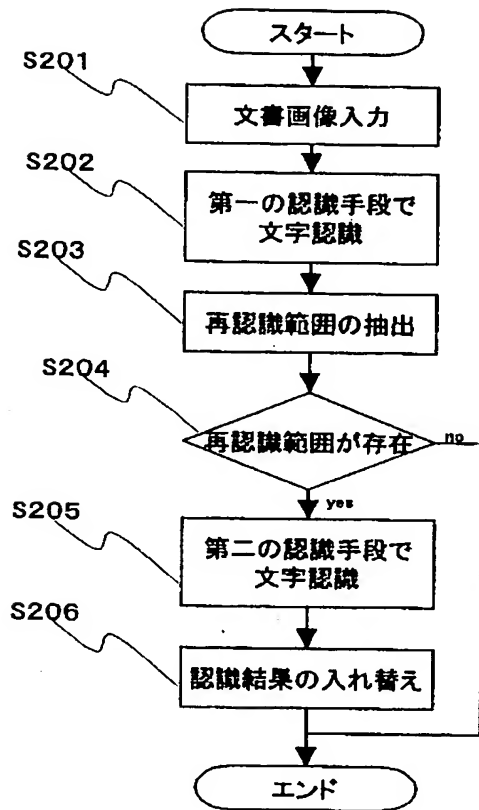
【書類名】

図面

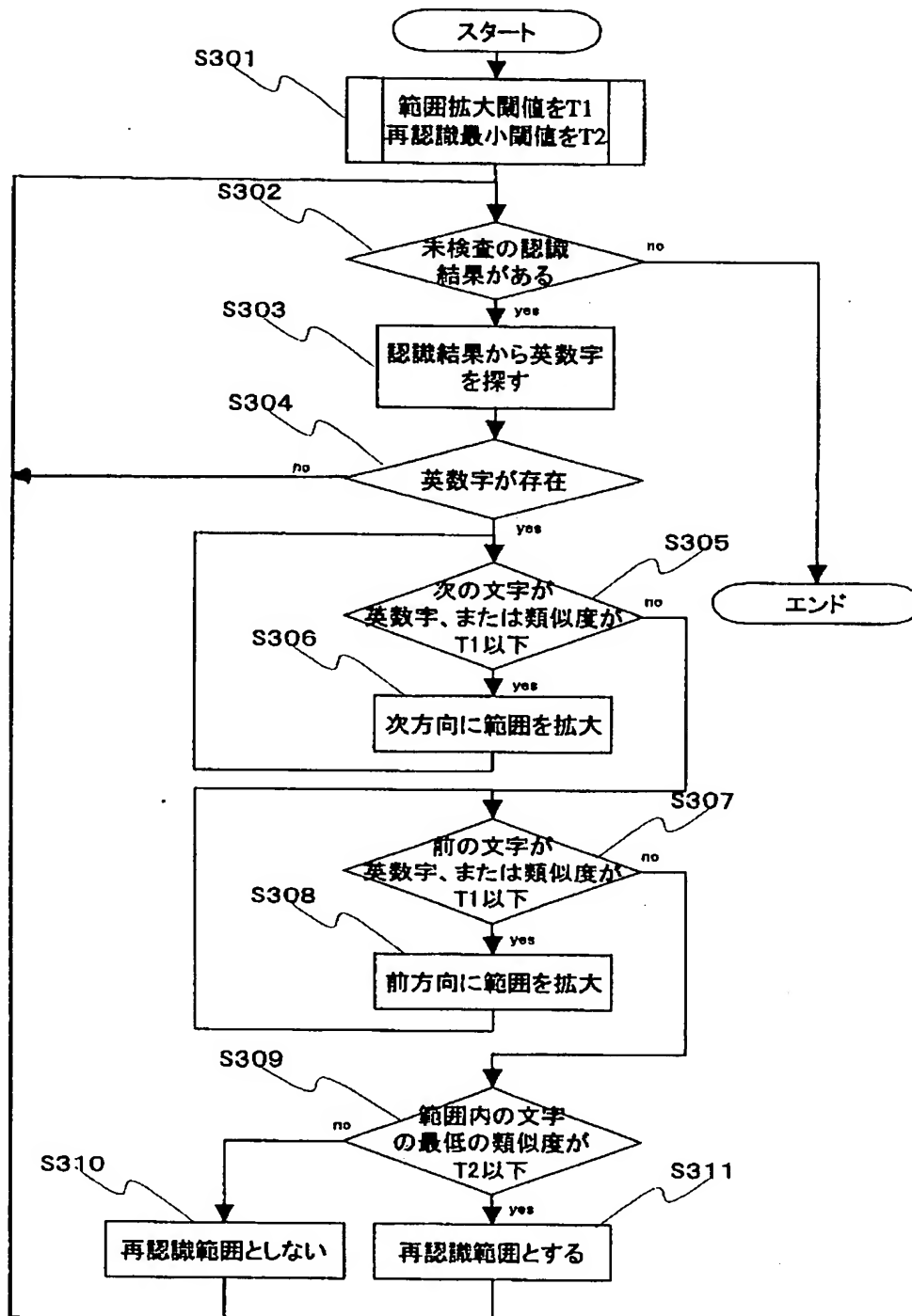
【図 1】



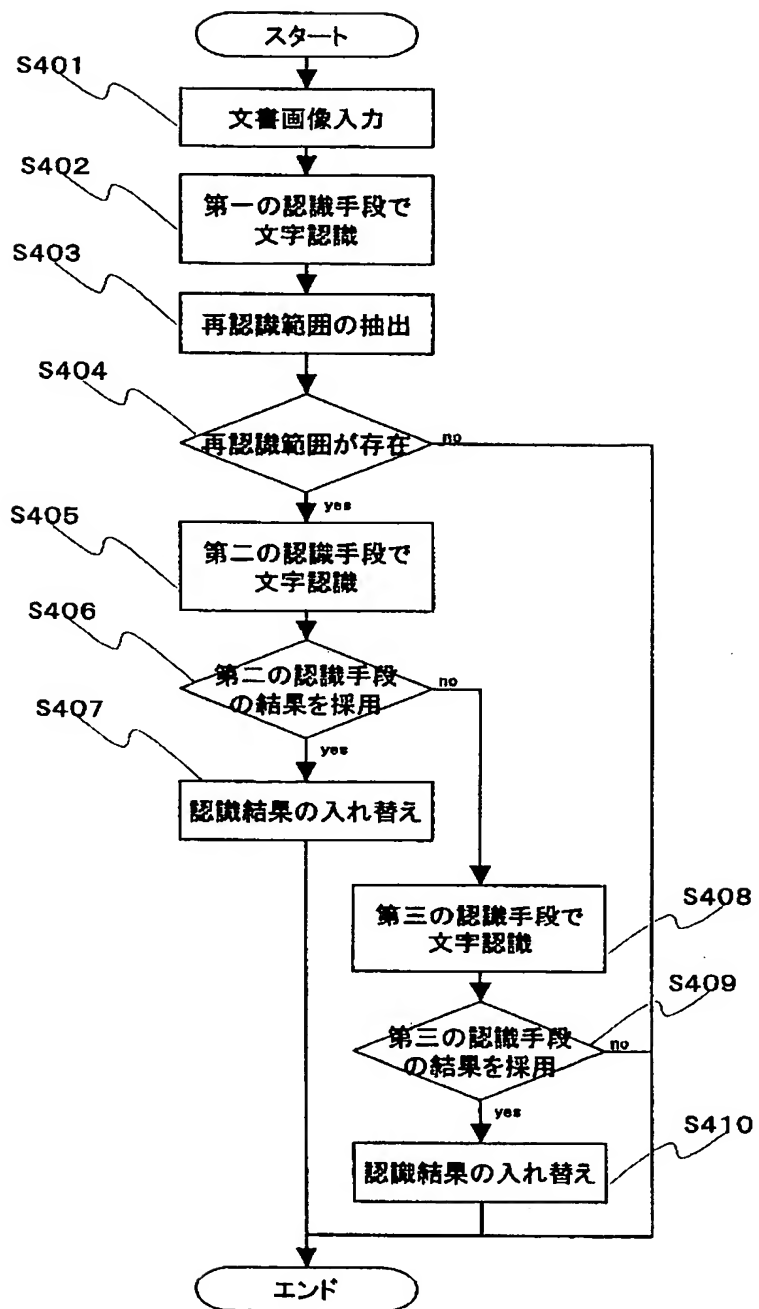
【図 2】



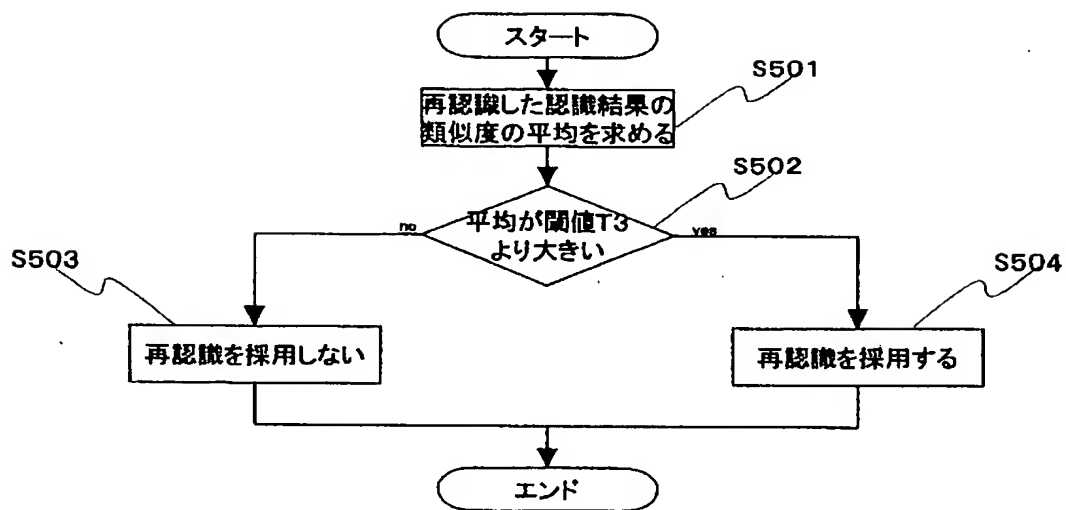
【図 3】



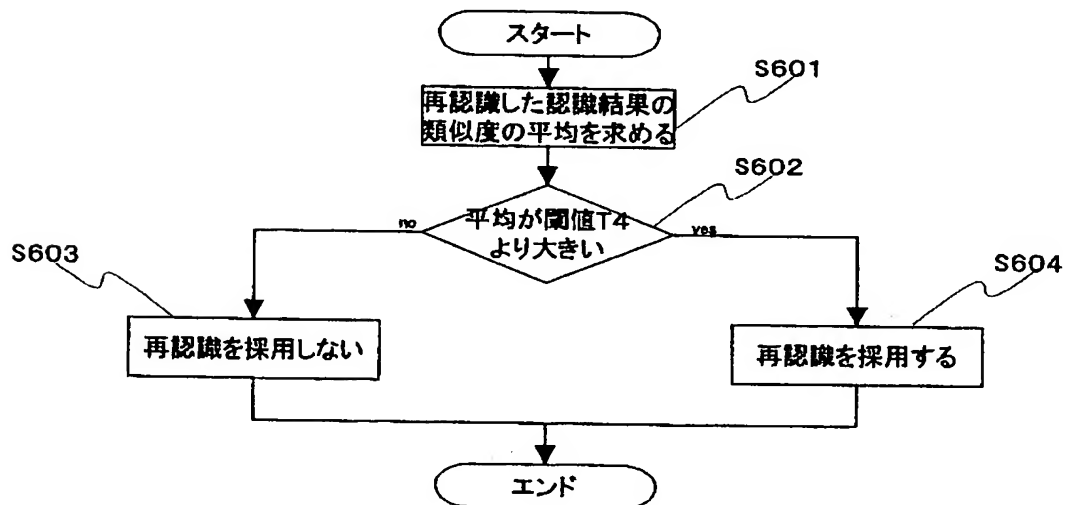
【図 4】



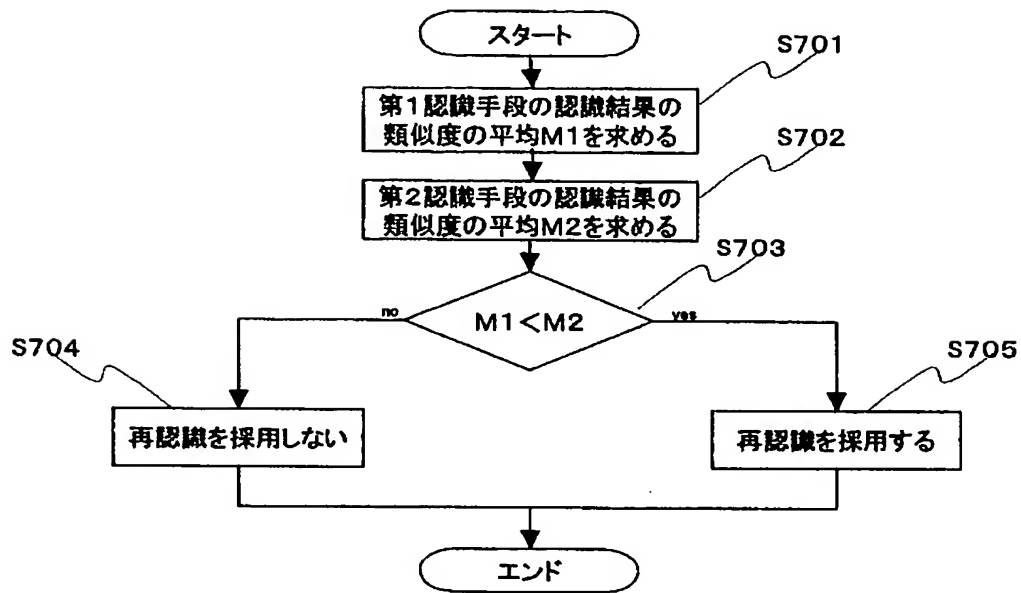
【図 5】



【図 6】



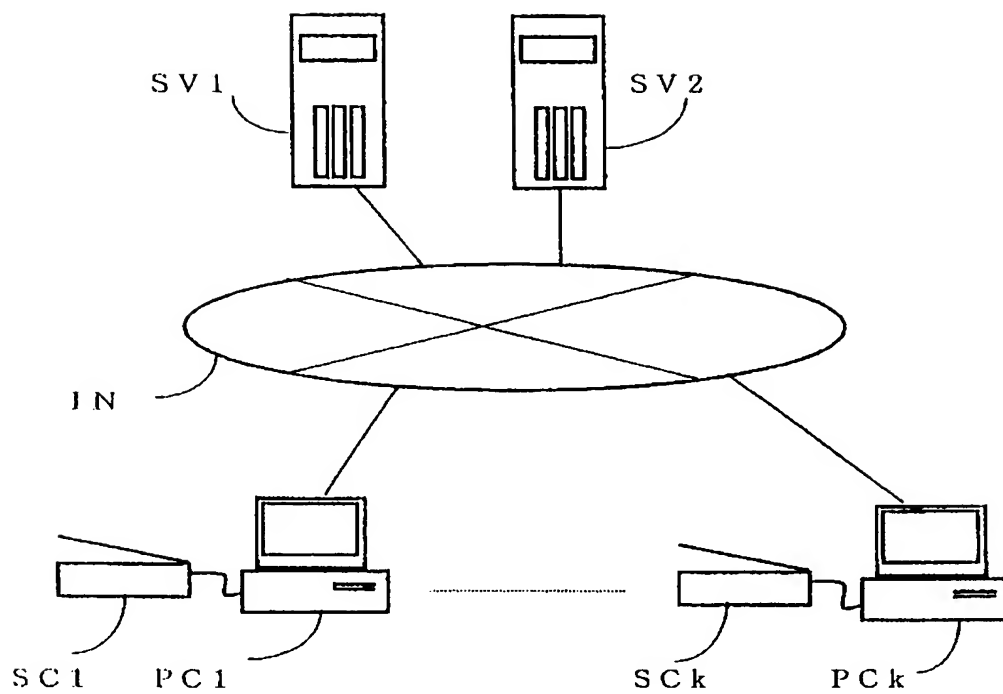
【図 7】



【図 8】

ディレクトリ情報	A 0
第 1 のデータ処理プログラム 図 2 に示すフローチャートのステップに 対応するプログラムコード群	A 1
第 2 のデータ処理プログラム 図 3 に示すフローチャートのステップに 対応するプログラムコード群	A 2
第 3 のデータ処理プログラム 図 4 に示すフローチャートのステップに 対応するプログラムコード群	A 3
第 4 のデータ処理プログラム 図 5 に示すフローチャートのステップに 対応するプログラムコード群	A 4
第 5 のデータ処理プログラム 図 6 に示すフローチャートのステップに 対応するプログラムコード群	A 5
第 6 のデータ処理プログラム 図 7 に示すフローチャートのステップに 対応するプログラムコード群	A 6
	A 7

【図 9】



【書類名】 要約書

【要約】

【課題】 複数の言語を含む文書画像の文字認識精度を高める。

【解決手段】 第1の認識手段を用いて認識した結果に基づいて、再認識範囲を設定し、第2の認識手段によって再度文字認識する。再認識した結果、第1の認識手段よりも高い類似度が得られたときは、その再認識範囲について、第1の認識手段の認識結果を、第2の認識手段の認識結果で置き換える。

【選択図】 図3

認定・付加情報

特許出願の番号	特願 2002-199505
受付番号	50201001224
書類名	特許願
担当官	第七担当上席 0096
作成日	平成 14 年 7 月 10 日

<認定情報・付加情報>

【特許出願人】

【識別番号】	000001007
【住所又は居所】	東京都大田区下丸子 3 丁目 30 番 2 号
【氏名又は名称】	キャノン株式会社

【代理人】

申請人

【識別番号】	100079832
【住所又は居所】	東京都千代田区鍛冶町 1-6-15 共同ビル（ 神田駅前）22 号 つくし特許事務所
【氏名又は名称】	山本 誠

【選任した代理人】

【識別番号】	100078846
【住所又は居所】	東京都千代田区鍛冶町 1 丁目 6 番 15 号 共同ビ ル（神田駅前）22 号 大音・田中特許事務所
【氏名又は名称】	大音 康毅

【選任した代理人】

【識別番号】	100087583
【住所又は居所】	東京都千代田区鍛冶町 1 丁目 6 番 15 号 共同ビ ル（神田駅前）22 号 大音・田中特許事務所
【氏名又は名称】	田中 増顕

次頁無

特願 2 0 0 2 - 1 9 9 5 0 5

出 願 人 履 歴 情 報

識別番号

[0 0 0 0 0 1 0 0 7]

1. 変更年月日

1 9 9 0 年 8 月 3 0 日

[変更理由]

新規登録

住 所

東京都大田区下丸子 3 丁目 3 0 番 2 号

氏 名

キャノン株式会社